

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/45275 A2

(51) International Patent Classification⁷: **H04B**

(21) International Application Number: **PCT/IL01/01095**

(22) International Filing Date: 28 November 2001 (28.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/253,087 28 November 2000 (28.11.2000) US

(71) Applicant (for all designated States except US): **FLASH NETWORKS LTD.** [IL/IL]; Maskit street 15, 46733 Herzliya (IL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SHAHAM, Oded** [IL/IL]; Kiryat Sefer Street 20, 34676 Haifa (IL). **SAGIV, Avi** [IL/IL]; Anatot Street 34, 69080 Tel Aviv (IL). **SHAPIRA, Yair** [IL/IL]; D'Israeli Street 27, 34333 Haifa (IL). **ZAKAI, Noam** [IL/IL]; Pinski Street 13, 34351 Haifa (IL).

(74) Agents: **EITAN, PEARL, LATZER & CO-HEN-ZEDEK** et al.; Gav Yam Center 2, Shenkar Street 7, 46725 Herzlia (IL).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

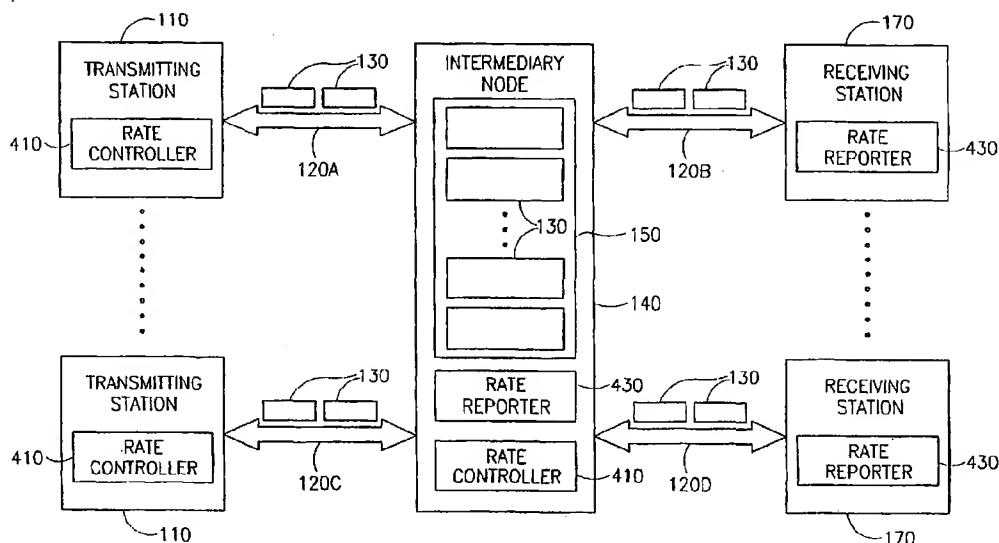
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR A TRANSMISSION RATE CONTROLLER



(57) Abstract: A method for proposing at least one transmission rate change including calculating a plurality of latency values, computing at least one derivative-based proposed change from the plurality of latency values, and proposing a rate change selected from the at least one derivative-based proposed change. Also provided is a system including a rate controller controlling the transmission rate of data between two stations over a network and a rate reporter in communication with the rate controller.

WO 02/45275 A2

SYSTEM AND METHOD FOR A TRANSMISSION RATE CONTROLLER

FIELD OF THE INVENTION

The present invention relates generally to network based communication,
5 specifically to improvements in bandwidth use.

BACKGROUND OF THE INVENTION

Data communication networks are built from various communication lines connected to each other. Every communication line has a specific capacity (or bandwidth), generally measured in units of data per time, typically kilobits per second (kbps). When 10 data is transmitted over a communication line by some station in the network, hereinbelow referred to as the transmitting station, it is transmitted at the specific transmission rate of that communication line. Between the transmitting station and the communication line there is usually a queue of limited size, in which data waits to be sent if the line is occupied. If a station attempts to transmit data faster than the capacity of the line, the 15 excess data is stored in the queue until it can be sent. If too much data is sent to the queue, the buffer will overflow and data will be lost (or dropped). Therefore, a transmitting station has to adjust its transmission rate to the capacity of the line.

Furthermore, this queue may be distributed in several locations along the communication lines in buffers. Some of these buffers may be "far" from the sender and 20 may be located in intermediary nodes along the network. These buffers of intermediary nodes may not be controlled or measured by the sender. Thus, in adjusting its transmission rate, a transmitting station also must take into account the state of the queue (e.g. whether the queue is accumulating or emptying).

These networks use standard communication protocols to allow communication 25 between computers, for example, Transmission Control Protocol/Internet Protocol (TCP/IP) and User Datagram Protocol/Internet Protocol (UDP/IP). In the transfer of data or information between computers, standard methods are used, for example HyperText Transfer Protocol (HTTP) Internet browsers and HTTP servers, file transfer protocol (FTP) servers, etc. Unfortunately, the rate control mechanism standards known in the art are

generally loss based. For example, in TCP/IP communications the rate of transmission is increased until data loss is detected, guaranteeing lost packets that will have to be resent.

Users of data communication networks often experience severe communication constraints due to non-optimal handling of the data transmission. This may be due to an 5 inefficient use of the network bandwidth and a rate of data transmittal slower than necessary. Contributing to these are the inherent inefficiencies of existing communication protocol standards and the current explosion of network traffic, which is overloading the capacity of networks.

10

SUMMARY OF THE INVENTION

There is provided, in accordance with an embodiment of the present invention, a method for controlling the transmission rate of data between stations. The method may include calculating a latency value per packet of a plurality of packets and generating at least one rate change from the one latency value.

15

Moreover, in accordance with an embodiment of the present invention, the calculating uses times from two independent clocks.

Furthermore, in accordance with an embodiment of the present invention, the latency value may be a two-way latency or a one-way latency.

20

Still further, in accordance with an embodiment of the present invention, the method further includes calculating a plurality of discrete derivatives of the latency value.

Moreover, in accordance with an embodiment of the present invention the calculating further includes using a transmission rate of at least one packet from the plurality of packets.

25

Furthermore, in accordance with an embodiment of the present invention, the generating further includes using a sign-significance filter.

In addition, in accordance with an embodiment of the present invention, the generating further includes using a zero derivative increase.

Moreover, in accordance with an embodiment of the present invention, the generating further includes computing at least one derivative-based proposed change.

30

Still further, in accordance with an embodiment of the present invention, the generating further includes using a loss-based decrease.

Moreover, in accordance with an embodiment of the present invention, the generating further includes taking a minimum of the loss-based decrease and the derivative-based proposed change.

5 Furthermore, in accordance with an embodiment of the present invention, the generating further includes using a future rate change.

In addition, in accordance with an embodiment of the present invention, the generating further includes computing a weighted average of the at least one rate change.

Still further, in accordance with an embodiment of the present invention, the generating further includes using the capacity of at least one communication line.

10 Moreover, in accordance with an embodiment of the present invention, the protocol may be a connectionless protocol or connection protocol.

Furthermore, in accordance with an embodiment of the present invention, the connectionless protocol may be UDP/IP and the connection protocol may be TCP/IP.

15 Moreover, in accordance with an embodiment of the present invention, the generating may be able to be done on a transmitting side or on a reporting side.

There is also provided, in accordance with an embodiment of the present invention, a method for proposing at least one transmission rate change. The method may include calculating a plurality of latency values, computing at least one derivative-based proposed change from the plurality of latency values, and proposing a rate change selected from the at least one derivative-based proposed change.

20 Moreover, in accordance with an embodiment of the present invention, the calculating further includes using a sign-significance filter.

Furthermore, in accordance with an embodiment of the present invention, the calculating further includes choosing at least one discrete derivative of the latency.

25 In addition, in accordance with an embodiment of the present invention, the computing further includes checking packet loss.

Still further, in accordance with an embodiment of the present invention, the computing further includes selecting a minimum between a loss-based decrease and the at least one derivative-based proposed change.

Moreover, in accordance with an embodiment of the present invention, the computing further includes reducing the at least one derivative-based proposed change by a future rate change.

5 Furthermore, in accordance with an embodiment of the present invention, the proposing further includes figuring a weighted average of results of the computing.

In addition, in accordance with an embodiment of the present invention, the proposing further includes comparing results of the computing to line capacity.

10 There is also provided, in accordance with an embodiment of the present invention, a system including a rate controller controlling the transmission rate of data between two stations over a network and a rate reporter in communication with the rate controller.

Moreover, in accordance with an embodiment of the present invention, the rate controller is located in a transmitting site.

15 Furthermore, in accordance with an embodiment of the present invention, the rate controller includes at least one transmitting packets table.

In addition, in accordance with an embodiment of the present invention, the rate reporter is located in a receiving site or at an intermediary node along the network.

Moreover, in accordance with an embodiment of the present invention, the rate reporter includes at least one report table.

20 Furthermore, in accordance with an embodiment of the present invention, the rate reporter may include a unit adapted to send the at least one report table to the rate controller.

25 In addition, in accordance with an embodiment of the present invention, the rate controller or the rate reporter may include a unit adapted to compute a proposed rate change.

Moreover, in accordance with an embodiment of the present invention, the system may include a unit adapted to compute a proposed rate change from a transmitting packets table.

30 Furthermore, in accordance with an embodiment of the present invention, the system may include a unit adapted to compute a proposed rate change from a report table.

In addition, in accordance with an embodiment of the present invention, the rate controller and the rate reporter are able to communicate with independent clocks.

There is also provided, in accordance with an embodiment of the present invention, a method for calculating a transmission rate change for a plurality of data units 5 using independent clocks on a transmission station and on a receiving station.

Moreover, in accordance with an embodiment of the present invention, the method includes comparing latency values of consecutive pairs of the plurality of data units.

10

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood and appreciated more fully from the following detailed description taken in conjunction with the appended drawings, in which:

15

Fig. 1 is a block diagram illustration of a communication environment comprising a transmission rate control system operative in accordance with an embodiment of the present invention;

Fig. 2 is a graphic illustration of exemplary changes in both the transmission rate the latency time of packets in the transmission rate control system of Fig. 1, operative in accordance with an embodiment of the present invention;

20

Fig. 3 is a block diagram illustration of the rate transmission control system of Fig. 1, operative in accordance with an embodiment of the present invention;

Fig. 4 is a flow chart illustration of a method embodied by the controller interface of Fig. 3, operative in accordance with an embodiment of the present invention;

25

Fig. 5 is a flow chart illustration of a rate recalculation method executed by the controller manager of Fig. 3, operative in accordance with an embodiment of the present invention;

Fig. 6 is a flow chart illustration of the method of analyzing a single window section of Fig. 5, operative in accordance with an embodiment of the present invention; and

Fig. 7 is a flow chart illustration of a method embodied by the report manager of Fig. 3, operative in accordance with an embodiment of the present invention.

30

DETAILED DESCRIPTION OF THE PRESENT INVENTION

The present invention is a system and method to control data transmission rates. The present invention may comprise an improved system and method for maximizing bandwidth use while minimizing data loss. The present invention may allow the use of 5 independent clocks on the receiving and transmission stations without the requiring the synchronization of the clocks. The transmission rate control system and method of the present invention may be used for connection based and connectionless communication.

Fig. 1, to which reference is now made, is a block diagram illustration of a communication environment, which may comprise a transmission rate control system that 10 is operative in accordance with an embodiment of the present invention. The communication system may comprise at least one transmitting station 110, at least one communication line 120, optional intermediary nodes 140, and at least one receiving station 170. The rate controlled transmission system may be comprised of two units, a rate controller 410 and a rate reporter 430. Rate controller 410 may be located in transmitting 15 station 110. Rate reporter 430 may be located in receiving station 170. Intermediary nodes 140 may comprise a buffer 150, rate controller 410, and rate reporter 430.

Data may be sent from a given transmitting station 110, to a given receiving station 170 over at least one communication line 120A, 120B, 120C, or 120D. The data stream may be divided into chunks 130 of a predetermined size, which may be transmitted 20 over a series of communication lines, for example, 120A - 120D, using a known protocol. Along the path between the transmitting station 110 and the receiving station 170, there may be any number of intermediary nodes 140. These intermediary nodes 140 may receive the data and transmit it onwards until the data reaches its destination. Communication lines 120A, 120B, 120C, and 120D may be of the same type or different types. 25 Hereinbelow all communication lines 120A - 120D are referred to together as communication lines 120. Such a communication process using transmitting stations 110, communication lines 120, intermediary nodes 140, and receiving stations 170 and standards for its implementation are well known in the art.

The data may be transmitted at a certain rate, which may be measured in kilobits 30 per second (kbps). The transmission rate may be affected by the type of communication line 120 used, since different line types may have a different capacities or bandwidth. For

example, data may be divided into packets and may be sent over plain old telephone lines (POTS), Integrated Services Digital Network lines (ISDN), Terrestrial 1 data line (T1), etc. and may use UDP/IP, TCP/IP or any other protocol known in the art. For example, data may be divided into packets when a packet-based protocol such as TCP/IP is used. 5 Hereinbelow the term packets 130 will be used in place of chunks 130 due to the familiarity of the term. It is noted however, that any type of data division as defined in a communication standard may be used.

10 Intermediary node 140 may be any type of node along communication line 120, for example, a cell in cellular communication, a satellite station in satellite communication, a gateway between networks, etc. Any number of intermediary nodes 140 may be passed by the data during its transmission through the network. Any of the intermediary nodes 140 may function as a transmitting station 110 and/or a receiving station 170. If so, they may contain their own rate controller 410 and/or rate reporter 430. When acting in either 15 of these capacities they function as per the descriptions of those stations. For simplicity of the description, all intermediary nodes 140 are represented by a single element 140 in Fig. 1 and are referred to together.

20 Since the transmission rate from intermediary node 140 may be different than its receiving rate, there may be a need for buffer 150 to store a queue of the packets 130 that may be waiting to be sent. Buffer 150 may be under the sole control of intermediary node 140. Neither transmitting station 110 nor receiving station 170 may have access to buffer 150. Intermediary node 140 may transmit packets 130 to receiving station 170 over communication lines 120.

25 As will be explained in detail hereinbelow with respect to Fig. 3, transmitting station 110 may accumulate information about the transmissions that it sends. This may be done using its rate controller 410 and rate reporter 430 of the receiving station 170 to which the data was sent. Rate controller 410 may then modify the transmission rate of transmitting station 110 as will be explained hereinbelow with respect to Figs. 2 and 3. Two types of information may be collected by transmitting station 110, the latency of 30 packets and the packet losses. If the latency values are constantly increasing or constantly decreasing, it may indicate that the queue is constantly becoming longer or shorter, respectively.

Reference is now made to Fig. 2, a graphic illustration of exemplary changes in both the transmission rate of packets 130 and the latency time in a communication system comprising a transmission rate control system, operative in accordance with an embodiment of the present invention. Latency is a measure related to the time it takes for a 5 transmission to be received.

There are two types of latency, one-way latency and two-way latency. One-way latency is the time between the moment the data was transmitted by the transmitting station and the moment it was received by the receiving station. Two-way latency, also known as 10 round-trip time (RTT), is the time between the moment the data was transmitted by the transmitting station and the moment the transmitting station received an acknowledgment of the receipt of this data. This assumes that the acknowledgment was sent immediately 15 after the data was received.

The latency may comprise the following time intervals:

1. the waiting time in queues along the way and
- 15 2. the time required for the serialization and transmission of the data over the line.

In the case of two-way latency, there may be two further time intervals:

3. the waiting time of the acknowledgment in the queues along the return channel 20 and
4. the time required for the serialization and transmission of the acknowledgment over the return line.

It is assumed that the time intervals 2 and 4 may be constant most of the time, and 25 that interval 3 may not vary greatly for different acknowledgements, since the queue of the receiving side of transmission station 110 may be empty most of the time. Thus, a change in latency may be a good indicator of a change in time interval 1, which may reflect the waiting time in queues along the way. Changes in waiting time in the queue may imply 30 changes in the state of the queue, e.g. the queue size may be increasing or decreasing. Thus, monitoring the latency may provide information about the state of the queue.

Transmission rate curve 30 (solid line) may begin transmitting at the maximum rate possible and may decrease and increase over time. In response to transmission rate 35 curve 30, latency curve 20 (dashed line) may first rise at a dramatic rate as the system begins data transmission. As latency curve 20 increases, the transmission bit rate may be

lowered so that data packets 130 may not be lost. Over time, latency curve 20 may settle into a generally consistent latency time that may fluctuate slightly due to changing conditions over the communication lines 120 and the intermediate node 140.

At time T_0 , transmission of packets 130 may begin at a predetermined maximum rate (e.g. the maximum bit rate of the line). Latency curve 20 may rise until time T_1 , at which point the latency may be equal to a value predetermined by the method of the present invention. At this time, the transmission rate may be lowered as seen on transmission rate curve 30. Latency curve 20 may still rise, but its slope may change, reflecting the new transmission rate. At time T_2 , there is a sharp drop in latency curve 20, which may reflect a packet being lost. This may cause the method of the present invention to lower the transmission rate to a predetermined percent of line capacity, as reflected in transmission rate curve 30 at time T_3 .

The method of the present invention continues monitoring latency times and adjusts transmission times. From time $T_4 - T_5$, the latency time is almost constant, which may indicate that the transmission rate may be increased T_6 . Adjustments may be made in the transmission rate that may cause a change in latency time. The method may find a possibly optimal transmission rate for the current communication line conditions as explained hereinbelow, which may be used until a change in conditions is detected.

If the latency values are constantly increasing or constantly decreasing, it may imply that the queue is constantly becoming longer or shorter, respectively. In such a case, the difference between each two consecutive latency values may yield a correction to the current transmission rate. Such a measure may be referred to as the "discrete derivative of the latency" and its use may be referred to as correcting the transmission rate according to the discrete change in the latency time.

There may be situations in which a policy is required to govern dynamic changes to the transmission rate of transmitting station 110. Such a policy may be implemented by the transmission rate control system of the present invention. Examples of such situations may include an unknown line capacity, a line capacity that changes with time, a line that is shared by a varying number of transmitting stations, and a varying number of data flows that may originate from the same station ("shared medium") but may transmit at different rates.

Packet-switching cellular networks provide an example of the above-mentioned situations. The capacity dedicated to data users may change in accordance with the current number of voice users in the cell. The cell capacity is shared among a changing number of clients, which may each transmit at varying rates. Furthermore, a client may switch from a 5 cell having certain conditions (capacity, number of other clients) to a cell with different conditions.

Fig. 3, to which reference is now made, is a detailed block diagram illustration of the rate transmission control system of the present invention, which may comprise rate controller 410 and rate reporter 430 (of Fig. 1), operative in accordance with an 10 embodiment of the present invention. Also shown is network 420, which represents those communication lines 120 and intermediary nodes 140 that may occur between rate controller 410 and rate reporter 430. In the embodiment of Fig. 3 the parts of the rate transmission control system may be located in different locations – rate controller 410 in transmitting station 110 (Fig. 1) and rate reporter 430 in receiving station 170 (Fig. 1). In a 15 further embodiment of the present invention, they may be in the same location.

Whereas, rate controller 410 may control the transmission rate to more than one rate reporter 430, all calculations and rate adjustments may be done for each rate reporter 430 individually. It is also noted that a given rate reporter 430 may report to more than one 20 rate controller 410. It is noted that different report tables 460 may be generated for each connection and may be sent to the appropriate rate controller 410.

Rate controller 410 may comprise a controller manager 414 and a controller interface 413. Controller manager 414 may create at least one transmitting packets table 450 for each rate reporter 430 it controls. Upon receipt of data packet 130 from transmission station 110 via data bus 411, controller interface 413 may create an entry in 25 transmitting packets table 450, which may comprise the unique identification number (ID) of the packet and the transmission time according to its clock. Optionally, the transmission rate of the packet may also be included. Transmitting packets table 450 may comprise the collected information about the timing of packet 130. Data packet 130 may then be forwarded to network 420 by controller interface 413 for transmission to receiving station 30 170.

Rate reporter 430 may comprise a reporter manager 432 and a reporter interface 433. Reporter manager 432 may create at least one report table 460 for each rate controller 410 to which it reports. When a packet 130 is received from transmitting station 110, by reporter interface 433, reporter manager 432 may create an entry in report table 460, which 5 may be comprised of the ID of the packet and the receipt time according to its local clock. When a predetermined condition is met, for example a given number of packets have been received or a certain period of time has elapsed, reporter interface 433 may send report table 460 to rate controller 410 over network 420. Report table 460 may be sent as a packet according to any protocol known in the art, for example UDP/IP.

10 It is noted that if rate reporter 430 is located at receiving station 170, then the reported time may be the time the packet is received. If rate reporter 430 is located at intermediary node 140, it may be placed at the output end of the node, after the processing units and buffers. Thus the reported time from intermediary node 140 may be the time at which packet 130 is transmitted to its next destination.

15 When a report table 460 is received by controller interface 413 from rate reporter 430, controller manager 414 may associate it with the transmitting packets table 450 created for that rate reporter 430. Controller manager 414 compares the two tables. It may process the two relevant tables as described below. Based on the results of this process and the type of communication line 120, if line type information is available, controller 20 manager 414 may calculate a new recommended transmission rate and may transfer the new rate to transmitting station 110 via control bus 412.

25 Those skilled in the art will appreciate that the present invention may be implemented as software that may reside in the transmitting computer and/or in the reporting computer or that it may be comprised of a computer connected to the communication line. Furthermore those skilled in the art will appreciate that in case of two-way latency (RTT) the transmitting side and the reporting side may be located in the same station.

30 Fig. 4, to which reference is now made, is a flow chart illustration of a method implemented by controller interface 413 (Fig. 3), operative in accordance with an embodiment of the present invention. When a packet is transmitted, a new entry may be created, by the controller interface 413 in the relevant transmitting packets table 450 (step

510). This entry may comprise the packet ID, the transmission time, and the transmission rate. Packet 130 may then be transmitted to its destination.

Controller interface 413 may check if a report table 460 has been received (step 520). If not it may return to step 510 and may await receipt of the next packet. Report 5 table 460 may contain the receipt time for the packets handled by a given receiving station 170 (according to its clock time). If report table 460 has been received, the appropriate transmitting packets table 450 may be updated with the receipt time reported for each packet (according to its ID number) in report table 460 (step 530). If a packet is not listed in report table 460 but a packet that was transmitted after it does appear in the report, the 10 missing packet may be marked as a lost packet. This may assume that the packets arrive at receiving station 170 in the same sequence in which they were transmitted.

Controller interface 413 may check if enough data has been received in the updated transmitting packets table 450 to perform statistical computations (step 540). If not, controller interface 413 may return to step 510. Otherwise, controller interface 413 15 may transfer the updated transmitting packet table 450 to the controller manager 414 (step 550), which may begin computations. While computations are performed, interface 413 may return to step 510, which may allow controller interface 413 to continue data transmission and receipt.

The following definitions and table may be useful with respect to Figs. 5 - 7 20 hereinbelow.

Examination window: The part of the transmitting packets table that is examined by the rate recalculation mechanism.

Window Section: A chunk of the examination window in which all the packets were transmitted at the same rate. In one window a number of sections may exist, 25 indicating packets were transmitted at several different rates during the period covered by the window.

Window Section Rate: The rate at which the packets in the section were transmitted.

Latency Calculation: *one way latency = reporting time - transmitting time.*

Delta Value: The difference between the latency values of two packets that were transmitted consecutively.

Table 1: Updated Partial Transmitting Packets Table

Packet Number	Transmission Time (T)	Trans. Rate	Receipt time (R)	Latency (L = R - T)	Delta (D = L _{m+1} - L _m)
5	10.15	7000	22.00	11.85	
6	10.16	7000	22.02	11.86	0.01
7	10.17	7000	22.04	11.87	0.01
8	10.18	7000	22.05	11.87	0.00

Table 1 encompasses part of a window section of an updated transmitting packets table 450. The packet number column may comprise the unique packet IDs. The transmission time column may comprise the time, T, at which the packet was sent according to the local clock of the transmitting station. The transmission rate column may comprise the rate at which the packets were sent. Note in this partial table, since only one window section is shown, the rate is the same for all the packets. The receipt time column may comprise the time, R, the local time at the receiving station when the packet was received, as reported in the report table. In the case of one way latency, the latency column may comprise the difference between the transmission time and the receipt time, R - T. The delta column may comprise the delta value, D, comprising the difference between the latency of two consecutive rows, L_{m+1} - L_m, of two consecutively sent packets.

This list of delta values may reflect a change in the latency. As may be seen in Table 1, the delta values of different packets may factor out the clock time differences by the subtraction of latency values. This method may thus afford the flexibility of using two independent clocks without worrying about clock synchronization.

Reference is now made to Fig. 5, a flow chart illustration of the rate recalculation method that may be executed by controller manager 414 of rate controller 410 (Fig. 3), operative in accordance with an embodiment of the present invention. Controller manager 414 may receive the updated transmitting packets table 450 (step 601). It may calculate the one-way latency for each entry. The delta between the latency of the current entry and of the previous entry may also be calculated, and both may be entered into updated transmitting packets table 450 (step 602).

The examination window may be divided into one or more window sections, wherein each section may have the same transmission rate (step 603). A proposed rate change may be calculated for each window section (step 604). This calculation is described in detail hereinbelow with respect to Fig. 7. The final proposed rate change may 5 be an average of all the window sections weighted by the number of delta values in each section.

Based on the proposed rate changes, controller manager 414 may decide to change the transmission rate of transmitting station 110 (step 605) and may send these instructions over control bus 412. However, first a comparison may be made to check that 10 the new rate does not exceed the capacity of communication line 120. If it does, the rate may be set to the value of the capacity of the communication line. Thus, the capacity of the communication line may function as an upper bound for the proposed transmission rate.

As may be understood from the descriptions of Figs. 4 and 5, the tasks of rate controller 410 may be divided between two separate units, controller manager 414 and controller interface 413. Thus, those functions that take longer may be performed by controller manager 414, which may leave controller interface 413 to control the data receipt and transmission functions. It is noted however, that in a further embodiment these two logical units may be combined in a single unit. 15

Fig. 6, to which reference is now made, is a flow chart illustration of a method 20 that may be used in analyzing a single window section, operative in accordance with an embodiment of the present invention. This figure corresponds to a single iteration of step 604 of Fig. 5 and may be performed once per window section.

Controller manager 414 may apply a "sign-significance filter" to each of the delta values in the section (step 701). This test may be used to determine whether the value is either "significantly positive" or "significantly negative". For example, it may check if the 25 percentage of positive values in the delta list exceeds a certain threshold. A similar check may be done for negative values. If there is a significant value, the change (or derivative) may be calculated as the median of the majority group (whether positive or negative) and the other delta values may be disregarded (step 702). Otherwise, the derivative of the 30 window section may be set to "0" (step 703).

Thus, the "sign-significance filter" may be used to determine whether there is a clear tendency in the delta values of the window section. This may be necessary as the use of the median may provide only a rough way to filter out exceptional values. Disregarding the other delta values may assist in a situation in which the tendency of the section's derivative is clear, yet exceptional values exist.

Controller manager 414 may calculate a single proposed rate change value for the window section from the derivative (step 704). This value may be called the "derivative-based proposed change". For a non-zero derivative, the derivative-based proposed change may be derived as follows. Given X milliseconds (msec) as the time between consecutive packet transmissions, and a derivative value of Y msec, the proposed rate change may be calculated so that the time between consecutive packet transmissions may be $X + Y$ msec.

A zero derivative may indicate one of the two following situations:

1. The transmission rate has reached a level in which it is adequate for the current conditions. The congestion level in the buffers is constant, and hence the derivative may be zero, or
2. The transmission rate is too low for the current conditions, but the buffers are totally empty. In such a case, the wait-time of all packets in the buffers may be "0" and therefore the derivative may be zero.

Since it may be hard to distinguish between the two situations and since under-utilization of the line may be considered more critical, a too low for transmission rate may be assumed (case 2). Thus, in the case of a derivative of zero the rate may be increased. Such an increase may be called a "zero derivative increase". This increase may be proportional to the capacity of the communication line (for example, 15% of the capacity). It is noted that even in the case of the first situation, the rate recalculations described hereinbelow may identify that the rate is too high, and may decrease it.

Controller manager 414 may next check if the percentage of lost packets, out of the total number of packets transmitted in the examined section, exceeds a certain threshold (step 705). If it does, the minimum between the derivative based proposed change and the "loss based decrease" may be used as the proposed rate change of the window section (step 706). The loss based decrease value may decrease the value by a

constant, which may be proportional to the line's capacity, for example, by 15 percent of capacity. If the derivative-based proposed change is not "negative enough" the rate may not be decreased enough and therefore the loss based decrease may be selected instead.

5 If the loss percentage does not exceed the threshold, the proposed rate change for the section may use the derivative based proposed change (step 707).

Since the packets in the examination window may have been transmitted prior to the calculation time, it may be the case that the transmission rate has changed. The current transmission rate may no longer be the same as at the time the section was transmitted. This fact may be taken in consideration when calculating the proposed rate change. Thus, 10 the proposed rate change may be reduced in relation to "future rate changes" (step 708). For example, if the rate of the window section was 9000 bit/sec, the proposed change – 3000 bit/sec, and the current transmission rate 7000 bit/sec, then the actual change needed is only –1000 bit/sec. This may be calculated by subtracting the current transmission rate from the transmission rate at the time the section was transmitted giving a rate difference. 15 Then this rate difference may be subtracted from the proposed rate change, giving a rate reflecting the current transmission value.

Reference is now made to Fig. 7, a flow chart illustration of a method executed by reporter manager 432 and reporter interface 433 of rate reporter 430 (Fig. 3), operative in accordance with an embodiment of the present invention. Reporter manager 432, per each 20 connection, may initialize a counter to zero and may create a report table 460 (step 801). Report table 460 may be used for recording the unique ID numbers of packets that are received and the time of receipt of the packet. Once a packet has been entered in report table 460 it may be referred to as a "reported packet". Reporter interface 433 may now wait for a packet to arrive (step 802).

25 Upon receipt of a packet, reporter interface 433 may update report table 460 as described and may increment the counter (step 803). Reporter interface 433 may then check whether the counter has reached a pre-defined value, which may be named "receive window size", for example, that the counter is equal to 100 (step 804). Other possible conditions may be that a pre-determined period of time has elapsed, or that a given 30 predetermined time has been passed. If not, reporter interface 433 may return to step 802 and may wait for the next packet.

Otherwise, reporter interface 433 may format report table 460 into a data packet and may add a transport protocol header to report table 460. For example, using UDP/IP, the header may include the address of transmitting station 110. Reporter interface 433 may then send report table 460 to rate controller manager 410 over network 420 (step 805), and 5 may return to initialization step 801. Note that report table 460 sent in step 805 may be used as input in step 520 (Fig. 4).

In a further embodiment of the present invention, the rate recalculations may be performed by report manager 432, in which case only the result may be reported to the transmitting side. This embodiment may require the following modifications to the 10 above-described system. Each packet sent by controller interface 413 may include its transmission time in its header. There may be a single table managed by reporter 430 instead of the two tables described hereinabove. When a packet arrives, both its transmission time and its arrival time may be available, and therefore the latency may be calculated and may be written into a new column in the table.

15 When there is enough information in the table, the rate recalculation may be performed by reporter manager 432 in a manner similar to that performed by controller manager 414 as described hereinabove with respect to Figs. 5 and 6. The new rate may be transmitted by reporter interface 433 to rate controller manager 410, which may update the transmission rate. It is noted that in this embodiment, transmission of report table 460 as 20 described hereinabove in step 805 may be unnecessary.

As may be understood from the descriptions above, the tasks of rate reporter 430 may be divided between two separate units, reporter manager 432 and reporter interface 433. It is noted however, that in a further embodiment of the present invention these two logical units may be combined in a single unit.

25 It is hereby noted that, within the description hereinabove, the word "packet" is merely a name for a chunk of data and that other names, such as "segment", "frame", etc., are also possible and are included within the scope of the present invention.

Alternative embodiments will become apparent to those skilled in the art to which the present invention pertains without departing from its spirit and scope. It will thus be 30 appreciated by persons skilled in the art that the present invention is not limited by what

has been particularly shown and described herein above. Rather the scope of the invention is defined by the claims that follow:

CLAIMS

What is claimed is:

1. A method for controlling the transmission rate of data between stations, the method comprising:
 - calculating a latency value per packet of a plurality of packets; and
 - generating at least one rate change from said latency value.
2. A method according to claim 1, wherein said calculating uses times from two independent clocks.
3. A method according to claim 1, wherein said latency value is a two-way latency.
4. A method according to claim 1, wherein said latency value is a one-way latency.
5. A method according to claim 1, and further comprising calculating a plurality of discrete derivatives of said latency value.
6. A method according to claim 1, wherein said calculating further comprises using a transmission rate of at least one packet from said plurality of packets.
7. A method according to claim 1, wherein said generating further comprises using a sign-significance filter.
8. A method according to claim 1, wherein said generating further comprises using a zero derivative increase.
9. A method according to claim 1, wherein said generating further comprises computing at least one derivative-based proposed change.
10. A method according to claim 9, wherein said generating further comprises using a loss-based decrease.
11. A method according to claim 10, wherein said generating further comprises taking a minimum of said loss-based decrease and said at least one derivative-based proposed change.

12. A method according to claim 1, wherein said generating further comprises using a future rate change.
13. A method according to claim 1, wherein said generating further comprises computing a weighted average of said at least one rate change.
14. A method according to claim 1, wherein said generating further comprises using a capacity of at least one communication line.
15. A method according to claim 1, wherein said protocol is a connectionless protocol.
16. A method according to claim 15, wherein said connectionless protocol is UDP/IP.
17. A method according to claim 1, wherein said protocol is a connection protocol.
18. A method according to claim 17, wherein said connection protocol is TCP/IP.
19. A method according to claim 1, wherein said generating is able to be done on a transmitting side.
20. A method according to claim 1, wherein said generating is able to be done on a reporting side.
21. A method for proposing at least one transmission rate change comprising:
 - calculating a plurality of latency values;
 - computing at least one derivative-based proposed change from said plurality of latency values; and
 - proposing a rate change selected from said at least one derivative-based proposed change.
22. A method according to claim 21, wherein said calculating further comprises using a sign-significance filter.
23. A method according to claim 21, wherein said calculating further comprises choosing at least one discrete derivative of the latency.
24. A method according to claim 21, wherein said computing further comprises checking packet loss.

25. A method according to claim 21, wherein said computing further comprises selecting a minimum between a loss-based decrease and said at least one derivative-based proposed change.
26. A method according to claim 21, wherein said computing further comprises reducing said at least one derivative-based proposed change by a future rate change.
27. A method according to claim 21, wherein said proposing further comprises figuring a weighted average of results of said computing.
28. A method according to claim 21, wherein said proposing further comprises comparing results of said computing to line capacity.
29. A system comprising:
 - a rate controller controlling the transmission rate of data between two stations over a network; and
 - a rate reporter in communication with said rate controller.
30. A system according to claim 29, wherein said rate controller is located in a transmitting site.
31. A system according to claim 29, wherein said rate controller comprises at least one transmitting packets table.
32. A system according to claim 29, wherein said rate reporter is located in a receiving site.
33. A system according to claim 29, wherein said rate reporter is located at an intermediary node along said network.
34. A system according to claim 29, wherein said rate reporter comprises at least one report table.
35. A system according to claim 29, wherein said rate reporter comprises a unit adapted to send said at least one report table to said rate controller.
36. A system according to claim 29, wherein said rate controller comprises a unit adapted to compute a proposed rate change.

37. A system according to claim 29, wherein said rate reporter comprises a unit adapted to compute a proposed rate change.
38. A system according to claim 29, comprises a unit adapted to compute a proposed rate change from a transmitting packets table.
39. A system according to claim 29, comprises a unit adapted to compute a proposed rate change from a report table.
40. A system according to claim 29, wherein said rate controller and said rate reporter are able to communicate with independent clocks.
41. A method for calculating a transmission rate change for a plurality of data units using independent clocks on a transmission station and on a receiving station.
42. A method according to claim 41 comprising comparing latency values of consecutive pairs of said plurality of data units.

1/7

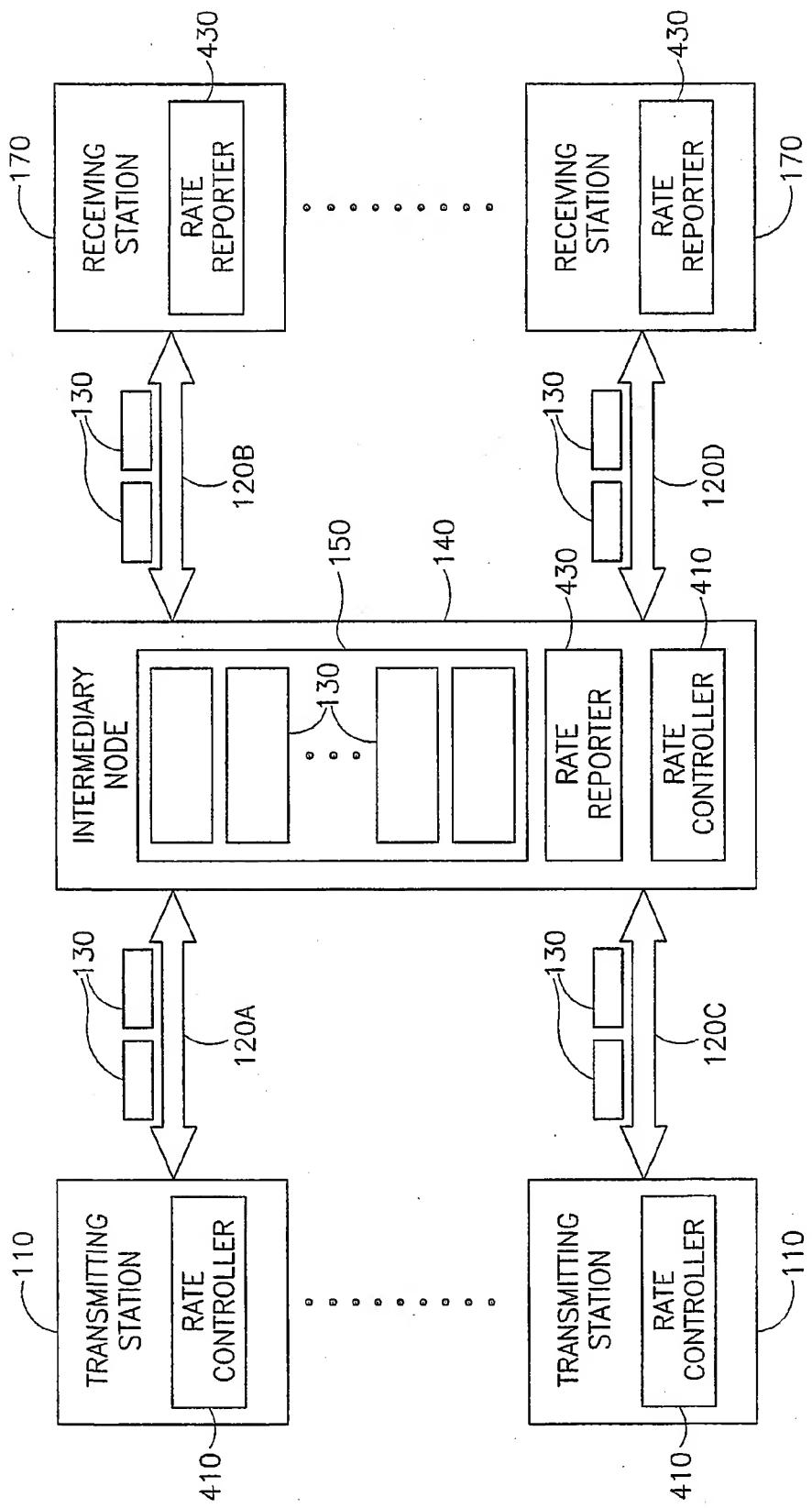
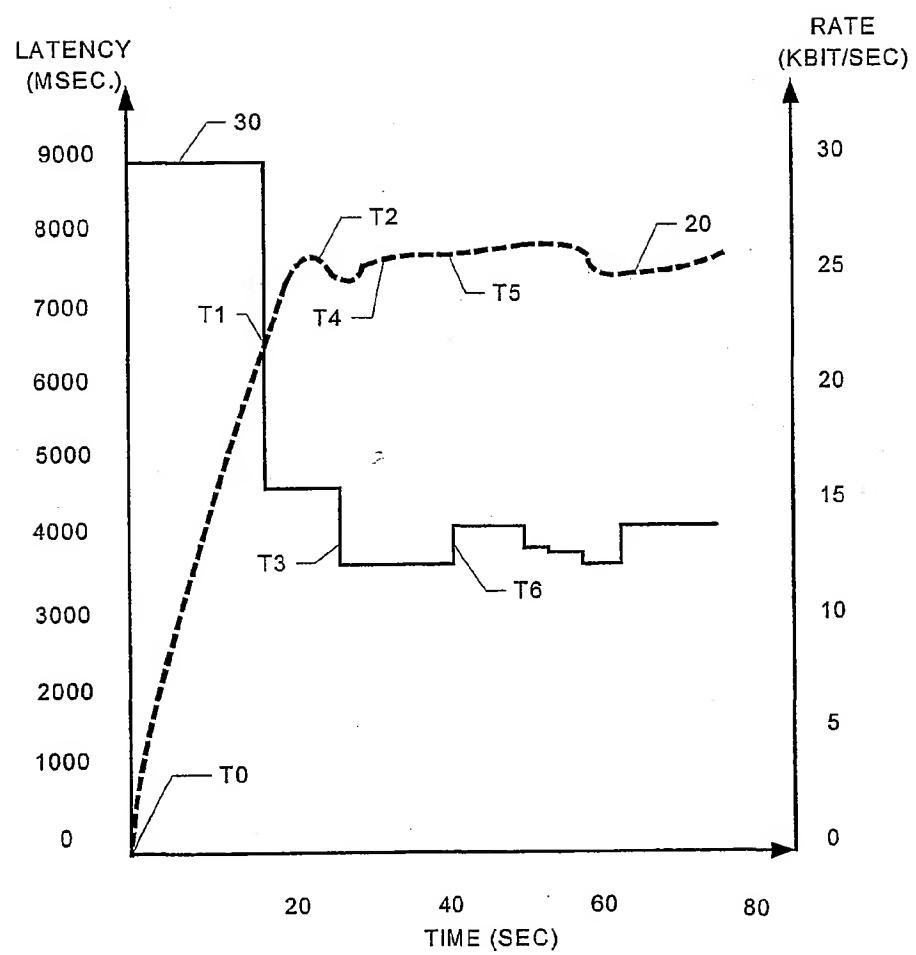


FIG.1

2/7

Fig. 2

3/7

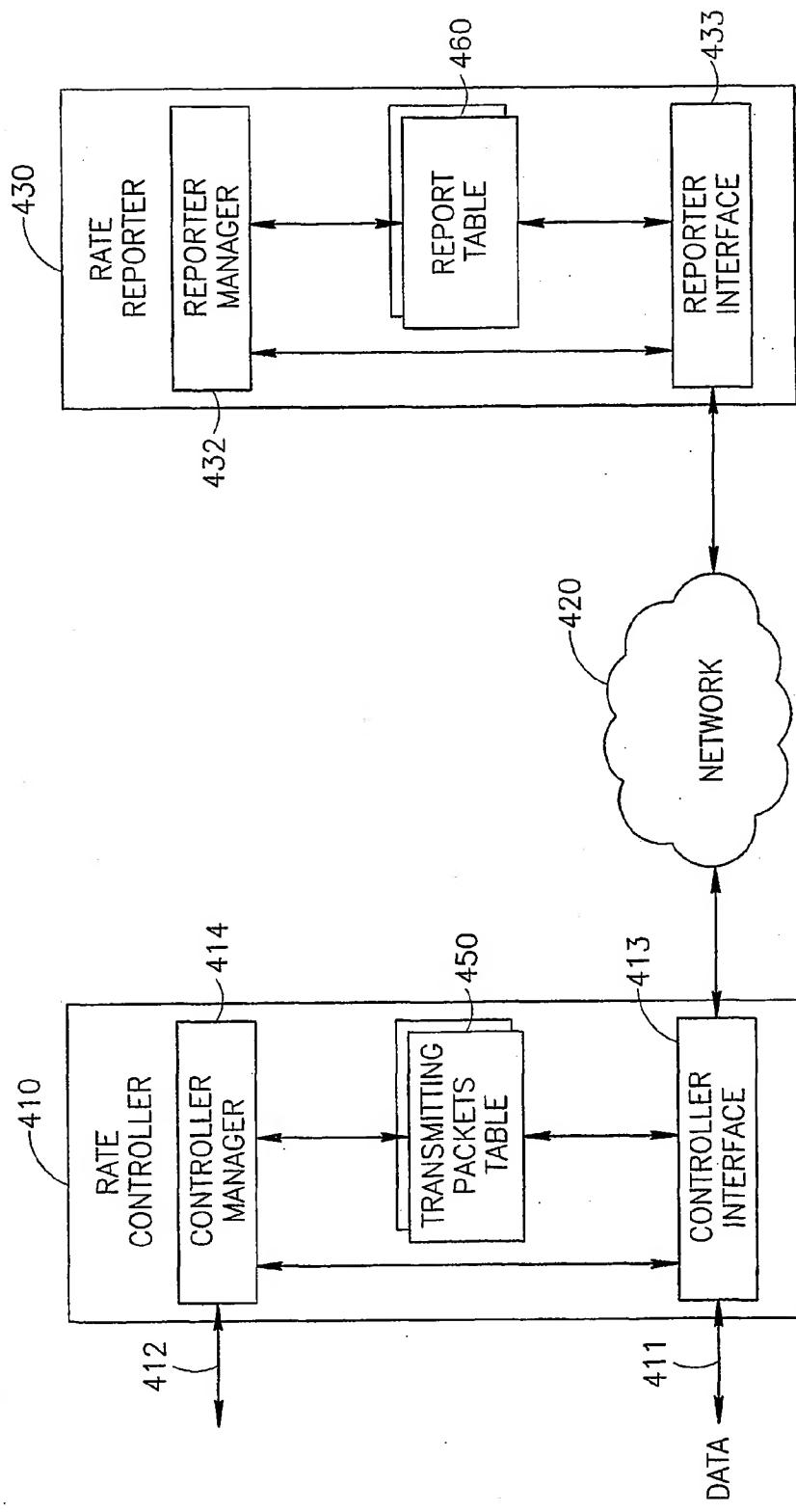
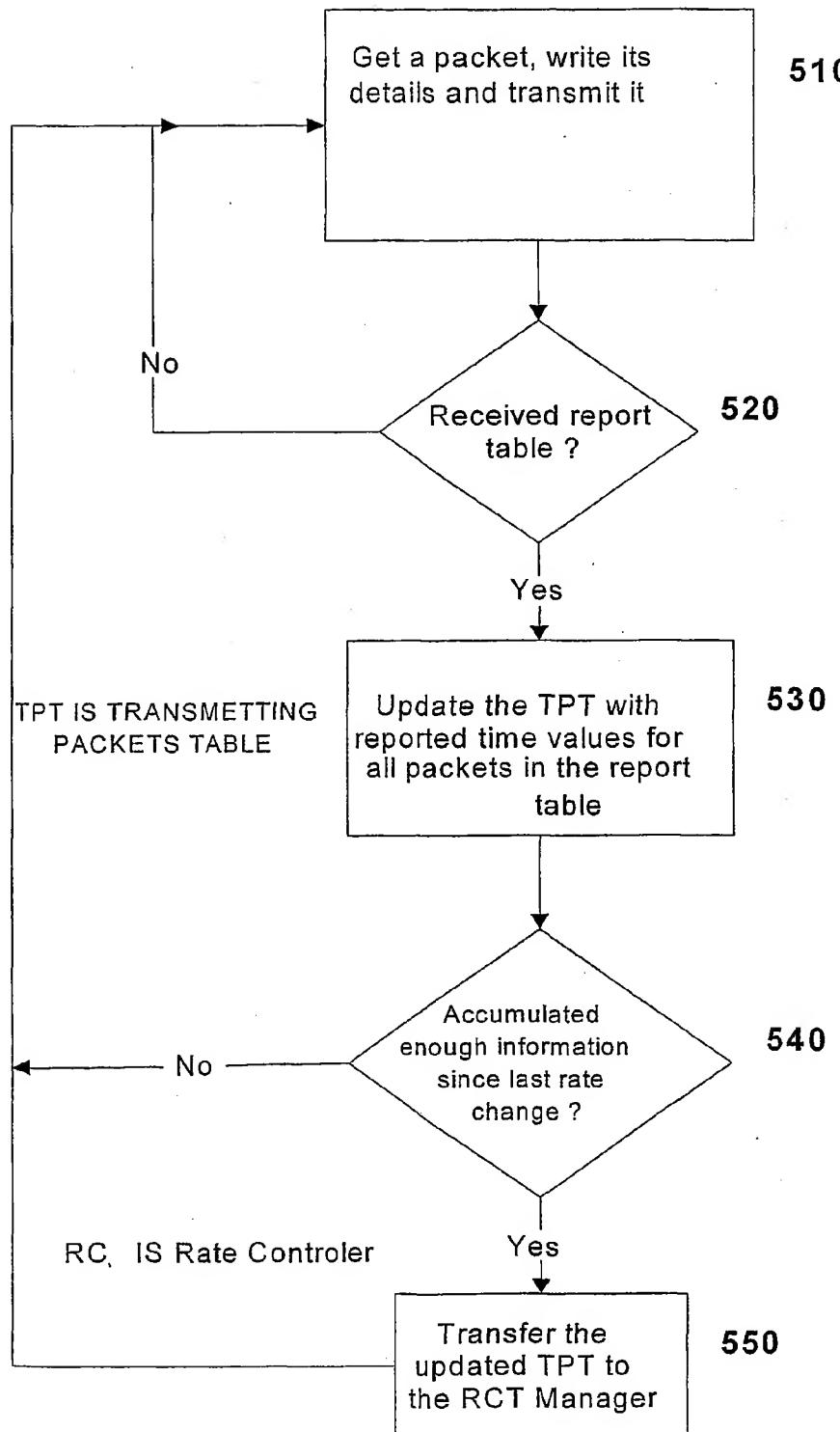


FIG.3

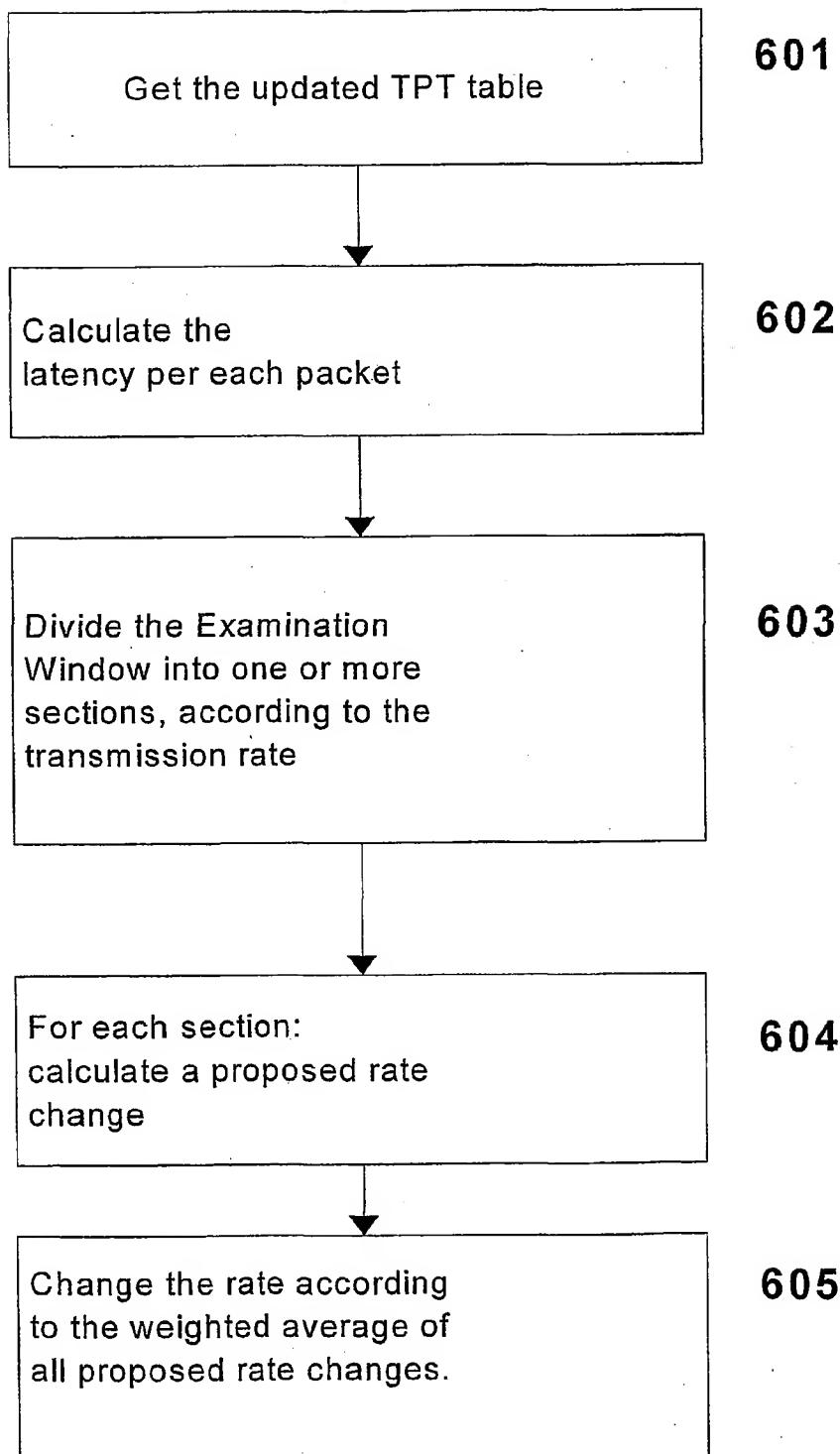
4/7

Fig. 4



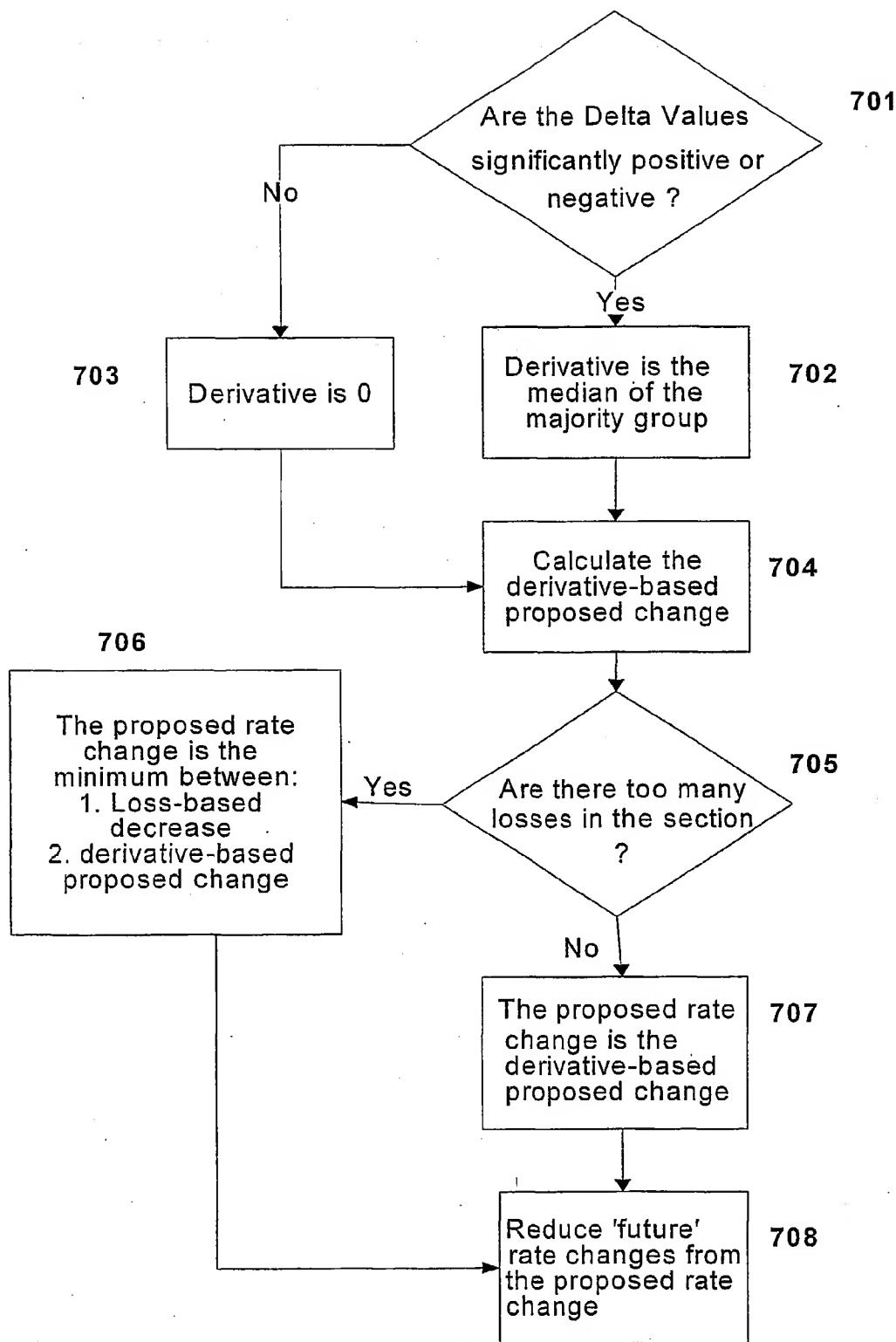
5/7

Fig. 5



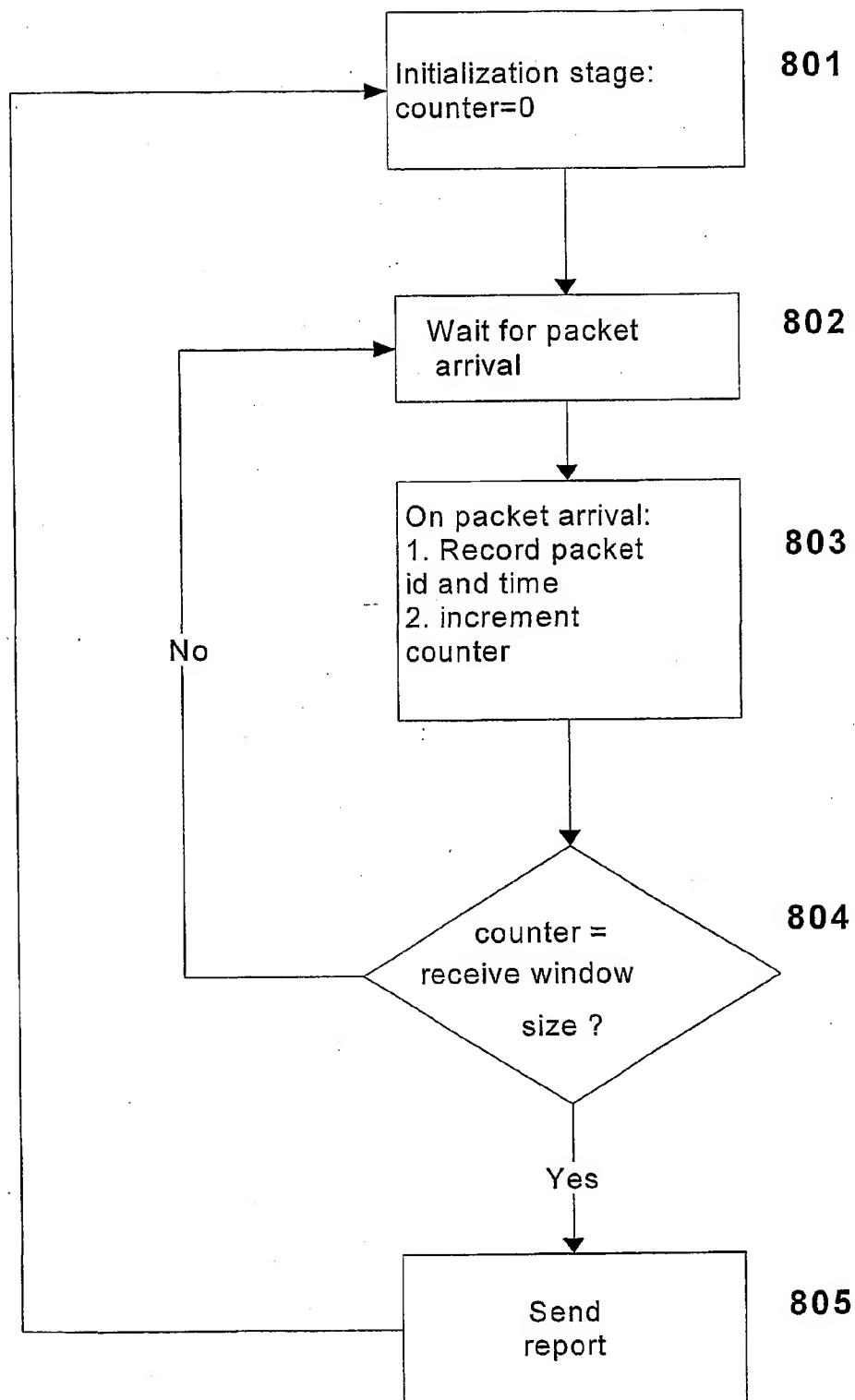
6/7

Fig. 6



7/7

Fig. 7



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/045275 A3

(51) International Patent Classification⁷: H04L 12/26 (74) Agents: EITAN, PEARL, LATZER & CO-HEN-ZEDEK et al.; Gav Yam Center 2, Shenkar Street 7, 46725 Herzlia (IL).

(21) International Application Number: PCT/IL01/01095 (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 28 November 2001 (28.11.2001) (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English (26) Publication Language: English (30) Priority Data: 60/253,087 28 November 2000 (28.11.2000) US (71) Applicant (for all designated States except US): FLASH NETWORKS LTD. [IL/IL]; Maskit street 15, 46733 Herzliya (IL).

(72) Inventors; and (75) Inventors/Applicants (for US only): SHAHAM, Oded [IL/IL]; Kiryat Sefer Street 20, 34676 Haifa (IL). SA-GIV, Avi [IL/IL]; Anatot Street 34, 69080 Tel Aviv (IL). SHAPIRA, Yair [IL/IL]; D'israeli Street 27, 34333 Haifa (IL). ZAKAI, Noam [IL/IL]; Pinski Street 13, 34351 Haifa (IL).

Published:

— with international search report

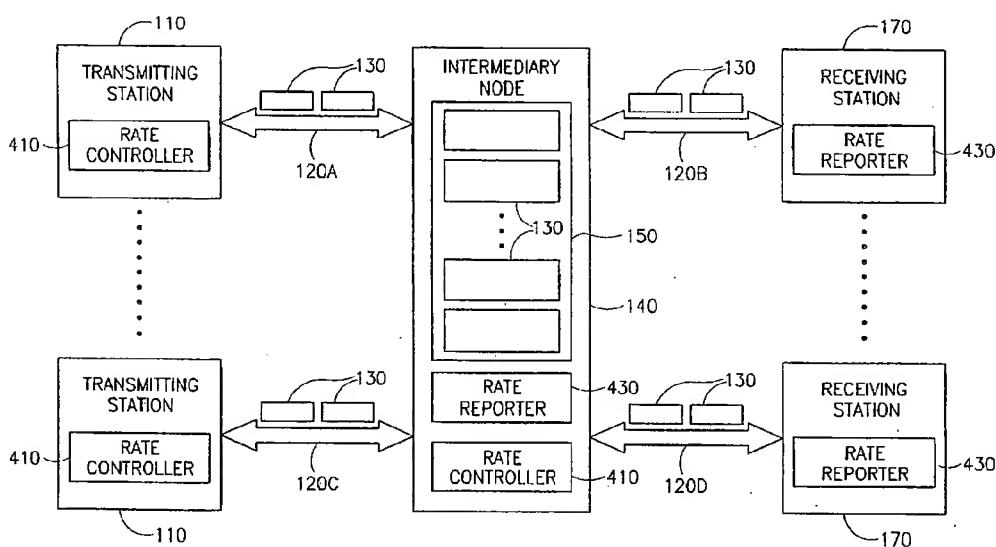
(88) Date of publication of the international search report: 12 September 2002

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR A TRANSMISSION RATE CONTROLLER



WO 02/045275 A3



(57) Abstract: A method for proposing at least one transmission rate change including calculating a plurality of latency values, computing at least one derivative-based proposed change from the plurality of latency values, and proposing a rate change selected from the at least one derivative-based proposed change. Also provided is a system (Fig. 1) including a rate controller (410) controlling the transmission rate of data between two stations over a network and a rate reporter (430) in communication with the rate controller (410).

WO 02/045275 A3



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL01/01095

A. CLASSIFICATION OF SUBJECT MATTER		
IPC(7) : H04L 12/26 US CL : 370/229-230.1,232-236,236.2,238,252-253,351-353,355-356,358,376,391,395, -395.21,395.5,395.52,395.61-395.62,395.64-395.65,397,399,409,428-429,468,477		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) U.S. : Please See Continuation Sheet		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Continuation Sheet		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,802,106 A (PACKER) 01 September 1998, figures 2-3	1-6,8-21,22-42
A	US 5,668,951 A (JAIN et al) 16 September 1997, figures 4-6	29-37
X	US 5,612,949 A (BENNETT) 18 March 1997, figures 1, 3-4, 7	1-6,8-21,23-28,40-42
X,P	US 6,205,120 B1 (PACKER et al) 20 March 2001, figures 2, 3A	1-6,8-21,23-29,36-37,40-42
A	US 5,014,265 A (HAHNE et al) 07 May 1991, figures 5-8	1-42
X	US 5,193,151 A (JAIN) 09 March 1993, column 2, lines 16-32	1,21,29,41
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		See patent family annex.
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 20 April 2002 (20.04.2002)	Date of mailing of the international search report 20 MAY 2002	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Phuongchau Ba Nguyen <i>Beljenia Zogar</i> Telephone No. 703-305-4700	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL01/01095

Continuation of B. FIELDS SEARCHED Item 1:

370/229-230.1,232-236,236.2,238,252-253,351-353,355-356,376,391,395.,-395.21,395.5,395.52,395.61-395.62,395.64-395.65,397,399,409,428-429,468,477

Continuation of B. FIELDS SEARCHED Item 3:

L1= (DYNAMIC\$4 AND (CONTROLS\$4 OR ADJUST\$4 OR ADAPT\$4 OR (INCREASE\$3 AND DESCRESS\$3))) SAME (RATE OR LOAD\$3)

L2= ((CALCULAT\$4 OR DETERMIN\$4 OR COMPUT\$4) SAME (LATENCY OR DELAY))